Roberta\_93

## Modèle de detection NLI pré-entrainé

Pour une version en anglais de notre API nous avons utilisé un modèle pré-entraîné issu de AllenNLP[1] entraîné sur les données SNLI qui est un ensemble de données seulement en anglais, représentant les légendes d’images, où ces dernières sont utilisées comme **prémisses**, les **hypothèses** sont créés manuellement par les travailleurs de Mechanical Turk.

1. Test de modèle pré-entrainé:

Données tests SNLI:

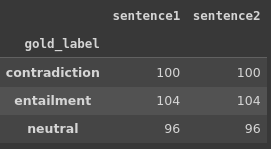


figure1: Échantillon de données test SNLI

Accuracy:

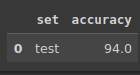
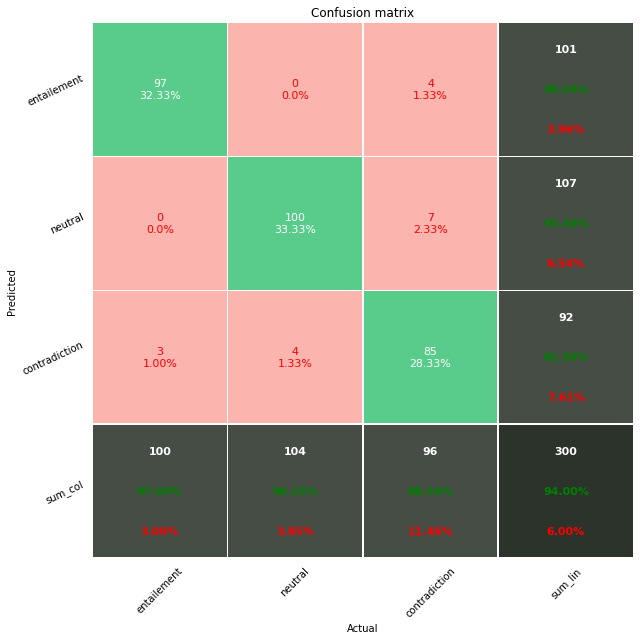


figure2:Precision de prediction sur les données SNLI test

Matrice de confusion:



Test de modèle sur les donnes contradictory:

Données tests contradictory:

Les données tests sont issues de jeu de données Contradictory, on a pris ce choix parce que ces données seront nouvelles pour le modèle pré-entrainé avec SNLI, cela nous permettra de tester vraiment la fiabilité de modèles et les performances enregistrés par ce modèle.

Les données tests sont comme suit:

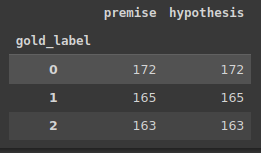


figure2:Données test

Accuracy:

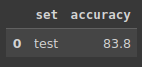
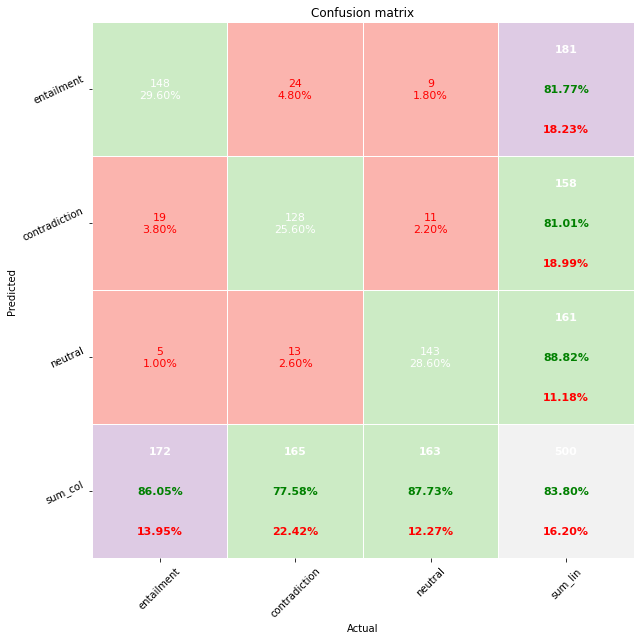


figure3: accuracy de données test contadictory



Discussion:

Nous remarquons une différence dans les performance de modèle sur les deux tests, tel que nous avons enregistré une baisse de performance de prédiction sur les données test contradictory, cela s’explique de fait que ces données sont nouvelles pour le modèles pré-entraîné, car se sont des données différentes de SNLi, rappelons que SNLI sont des données constituer à partir des légendes d’images, où ces dernières sont utilisées comme **prémisses**, les **hypothèses** sont créés manuellement par les travailleurs de Mechanical Turk.

De la nous constatons que l'issue(genre la manière dont elle se constituer ou leurs type comme snli c’est les legendes d’images, mais contradictory on sait pas d’ou elle proviennent ) des données influence sur la qualité de modèle.